Sleep Medicine 15 (2014) 530–535

Contents lists available at ScienceDirect

Sleep Medicine

journal homepage: www.elsevier.com/locate/sleep

Original Article

The structured Diagnostic Interview for Sleep Patterns and Disorders: rationale and initial evaluation

K.R. Merikangas [a,*], J. Zhang [a,1], H. Emsellem [b], S.A. Swanson [c], A. Vgontzas [d], F. Belouad [a], M.M. Blank [a], W. Chen [e], M. Einen [e], J.P. He [a], L. Heaton [a], E. Nakamura [a], S. Rooholamini [e], E. Mignot [e]

[a] Genetic Epidemiology Branch, Intramural Research Program, National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services, Bethesda, MD, United States
[b] The Center for Sleep and Wake Disorders, Chevy Chase, MD, United States
[c] Department of Epidemiology, Harvard School of Public Health, Boston, MA, United States
[d] Pennsylvania State College of Medicine, Hershey, PA, United States
[e] Department of Psychiatry and Behavioral Sciences, Stanford University, Palo Alto, CA, United States

ARTICLE INFO

Article history:
Received 9 May 2013
Received in revised form 10 September 2013
Accepted 8 October 2013
Available online 15 February 2014

Keywords:
Diagnostic Interview
Sleep disorders
Validity
International Classification of Sleep Disorders
Concordance
Area under the curve (AUC)

ABSTRACT

Objectives: We aimed to describe and report the initial validity of a newly developed structured interview for sleep disorders (Diagnostic Interview for Sleep Patterns and Disorders [DISP]) administered by trained lay interviewers.
Methods: A total of 225 patients with various sleep disorders were recruited from two nationally recognized sleep centers in the United States. The *International Classification of Sleep Disorders*, second edition (ICSD-2) criteria, were used to classify sleep disorders (e.g., delayed sleep phase disorder, hypersomnia, narcolepsy with cataplexy [NC], restless legs syndrome [RLS], periodic limb movement disorder [PLMD], insomnia, rapid eye movement sleep behavior disorder [RBD], and obstructive sleep apnea [OSA]). Interview diagnoses were compared with final diagnoses by sleep specialists (reference diagnosis based on clinical history, examination, and polysomnography [PSG] when indicated).
Results: DISP diagnoses had fair to substantial concordance with clinician diagnoses for various sleep disorders, with area under the receiver operator characteristic curves (AUC) ranging from 0.65 to 0.84. Participants classified by the clinician as having a sleep disorder were moderately well-detected (sensitivity ranging from 0.50 for RBD disorder to 0.87 for insomnia). Substantial specificity (>0.8) also was seen for five of the eight sleep disorders (i.e., delayed sleep phase, hypersomnia, NC, PLMD, and RBD). Interviews were more likely than clinicians to detect disorders secondary to the primary sleep problem.
Conclusions: The DISP provides an important tool for the detection of a wide range of sleep disorders in clinical settings and is particularly valuable in the detection of secondary disorders that were not the primary referral diagnosis.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

There has been increasing attention to the public health implications of sleep disorders. Insomnia [1–3], restless legs syndrome (RLS) [4,5] and obstructive sleep apnea (OSA) [6,7] are the most prevalent sleep disorders in the community and are associated with considerable daytime consequences [8] and negative health outcomes [9–11]. Because laboratory evaluation is required to establish a clinical diagnosis of several specific sleep disorders, including narcolepsy with cataplexy (NC), OSA, and rapid eye movement sleep behavior disorder (RBD) [12], it has been difficult to obtain accurate estimates of the magnitude of these disorders in the general population and their public health consequences. Some studies of community samples have utilized polysomnography (PSG) to determine the prevalence and mortality of specific disorders such as OSA [11,13], but the prohibitive cost and effort involved in administration of PSG is not feasible in large-scale community studies. Other population-based studies have solely focused on disorders that do not require PSG for a diagnosis such as insomnia or RLS [4,10,14–17]. Numerous questionnaires have been developed to collect information about specific sleep

* Corresponding author. Address: Intramural Research Program, National Institute of Mental Health, Genetic Epidemiology Research Branch, 35 Convent Dr, MSC# 3720, Bethesda, MD 20892, United States. Tel.: +1 301 496 1172; fax: +1 301 480 2915.
E-mail address: Kathleen.merikangas@nih.gov (K.R. Merikangas).
[1] Present address: Department of Psychiatry, Faculty of Medicine, The Chinese University of Hong Kong, China.

http://dx.doi.org/10.1016/j.sleep.2013.10.011
1389-9457/© 2014 Elsevier B.V. All rights reserved.

disorders, including insomnia [18,19], RLS [20–22], OSA [23,24], and general sleep quality [25]. However, most of these self-administered questionnaires do not assess the full criteria for these disorders. A third approach to estimate sleep morbidity statistics relies on identification of individuals in the general population who have received a diagnosis from a physician based either on laboratory or clinical evaluations. This approach can provide estimates of the full range of sleep disorders, but it is likely to yield underestimates, as only 33% of those with RLS in the general population have received a physician diagnosis [4].

Studies using one or more of these approaches have provided estimates of the magnitude of sleep disorders in community samples in specific regions of the United States and in international settings [13,14,26–40]. These studies estimate that approximately 30–48% of nonclinical samples report periodic insomnia [30], whereas approximately 10% report chronic insomnia (i.e., insomnia symptoms occurring at least three nights per week for 1 month) [30]. Population-based studies of RLS that have applied the *International Classification of Sleep Disorders*, second edition (ICSD-2), criteria have yielded prevalence estimates ranging from 5% to 12% [4,10,36], and those of OSA consistently yield prevalence estimates ranging between 2% and 7% of US adults based on either the apnea–hypopnea index or self-reports [6,41].

A major challenge to estimation of the magnitude of sleep problems and disorders in the general population is variation in the methods of assessing the diagnostic criteria for sleep disorders. Aside from the well-established self-report questionnaires for sleep difficulties, sleep patterns, and self-report assessments of specific sleep disorders including insomnia, OSA, RLS and NC, there is a lack of validated instruments for assessing ICSD-2 criteria for the full range of sleep disorders that can be administered by non-experts in the general population.

To address the lack of structured interviews for ICSD-2 sleep disorders, the National Institutes of Mental Health Intramural Research Program in collaboration with two sleep specialty clinics, the Stanford Center for Sleep Science and Medicine (Mignot E) and the Center for Sleep and Wake Disorders (Emsellem H) have developed a diagnostic interview called the Diagnostic Interview for Sleep Patterns and Disorders (DISP) to ascertain diagnostic criteria for the full range of sleep disorders. The interview covers sleep patterns, difficulties, and routines, as well as screens for most of the major sleep disorders. Respondents who screen into the key symptoms of a sleep disorder are then queried about the symptoms, course, impairment, severity, and treatment of each condition into which they positively screened. The DISP can be administered by lay interviewers either in person or by telephone. Computerized diagnostic algorithms have been developed to score the disorders according to ICSD-2 criteria. This interview is now included in the PHENX project of the National Institutes of Health (NIH) (https://www.phenx.org/). The goal of our report was to assess the concordance of the DISP with expert sleep clinician diagnoses for a range of sleep disorders in a clinical sample recruited from two US sleep centers.

2. Methods

2.1. Sample

The sample characteristics are presented in Table 1. A total of 225 subjects were recruited over a 2-year period from the Center for Sleep and Wake Disorders in Chevy Chase, Maryland (82.7%) and the Stanford Center for Sleep Science and Medicine in California (17.3%) for the interview validation study. The diagnoses included eight sleep disorders: delayed sleep phase disorder, hypersomnia, insomnia, NC, periodic limb movement disorder (PLMD), RLS, RBD, and OSA. The sample consisted of 57.3% women, with a mean age of 52.9 years (median, 54 years; range, 19–90 years). A majority of participants were non-Hispanic white, one-half married or cohabiting, and three-quarters with a college or higher-level education. Participants from Stanford Center were younger, more likely to be single and had lower education levels than those at the Chevy Chase Center. On average, administration of the interview was approximately 30 min.

2.2. Disorder assessment

2.2.1. The DISP

The DISP was developed by KRM and EM to assess both Sleep Patterns and Disorders in a structured format and it is to be used in the general population and clinical samples in nonsleep specialty settings. The interview is administered by trained lay interviewers with some medical background and experience in the administration of structured interviews. All interviewers received training from a clinical neurologist on the symptoms and diagnostic criteria of targeted sleep disorders, followed by clinical review of the interviews with feedback to establish common procedures, and joint blind ratings of interviews.

The first section of the interview includes questions regarding sleep patterns (i.e., routine sleep schedule, naps, sleep regularity, circadian preference). This section of the interview provides context regarding the general sleep-wake patterns and circadian preferences of the participants and a screen for potential delayed sleep phase disorder. The second section collects detailed information on the eight specific sleep disorders including 1–4 initial screening questions for each disorder followed by a more comprehensive evaluation of symptoms, signs, duration, course, and episodes (onset and offset history); impairment and severity; clinical examination; and help seeking and treatment histories if entry probes are endorsed. A module for advanced sleep phase disorder also has recently been added to the DISP. To avoid potential bias from using a clinical sample, the information about clinical examination, help seeking and treatment history was not included in the validation study. All diagnoses were based on lifetime symptoms and disorders.

The sleep disorders assessed in different sections of the DISP are the aforementioned eight sleep disorders. Computer algorithms were developed to assess criteria for these eight sleep disorders based on ICSD-2 criteria with minor modifications. For example, the classification for subtypes of insomnia was not further divided based on its assumed associated factors. For sleep disorders that require PSG examination or clinical examination in the ICSD-2 (NC and PLMD), the minimal criteria listed in ICSD, first edition, were used to determine the diagnoses [32]. For example, minimal criteria for NC (criterion B [recurrent excessive daytime sleepiness and naps] and criterion C [cataplexy]) were used to diagnose NC. Hypersomnia was defined as complaints of excessive daytime sleepiness symptoms occurring almost daily for at least 3 months. Multiple sleep latency test (MSLT) results (if any) were not used to diagnose hypersomnia in the computerized algorithms. The clinical interview and detailed algorithms for each disorder are available by request from the study author. The DISP is included on the NIH PHENX Toolkit for phenotype measures in genetic studies (https://www.phenx.org/) [42]. In addition, the key components of the diagnostic criteria for each sleep disorder in the DISP are presented in Table 2.

2.2.2. Clinician diagnoses

The patients in our study were adults referred to one of the two sleep centers. In suspected cases of hypersomnia or narcolepsy, patients underwent at least one night of standard PSG assessment and an additional MSLT. The clinician diagnosis was based on the disease history, clinical examination, PSG, and MSLT for all

Table 1
Demographic characteristics of sample from sleep centers ($n = 225$).

Characteristics	Total		Sleep center				Test statistic, P value
			Chevy Chase, MD		Stanford university		
	N	%	*n*	%	*n*	%	
Source	225	100.0	186	82.7	39	17.3	
Sex							$\chi^2_1 = 2.4$; $P = .121$
Women	129	57.3	111	59.7	18	46.2	
Men	96	42.7	75	40.3	21	53.8	
Marital status							
Never married	56	24.9	43	23.1	13	33.3	$\chi^2_3 = 41.1$; $P < .0001$
Married/cohabiting	111	49.3	102	54.8	9	23.1	
Separated/divorced/widowed	43	19.1	37	19.9	6	15.4	
Unknown	15	6.7	4	2.2	11	28.2	
Education							
High school or less	10	4.4	10	5.4	0	0.0	$\chi^2_1 = 18.4$; $P < .0001$ (college or higher vs others)
Some college	31	13.8	25	13.4	6	15.4	
College graduate	64	28.4	55	29.6	9	23.1	
Graduate/professional school	106	47.1	96	51.6	10	25.6	
Unknown	14	6.2	0	0.0	14	35.9	
Race/ethnicity							$\chi^2_1 = 0.1$; $P = .780$ (white vs other)
Hispanic	5	2.2	5	2.7	0	0.0	
Non-hispanic black	21	9.3	21	11.3	0	0.0	
Non-hispanic white	188	83.6	156	83.9	32	82.1	
Other including multiple race	4	1.8	2	1.1	2	5.1	
Age (y)	225	Mean = 52.9 (SD = 16.3); Min = 19, Max = 90	186	Mean = 54.1 (SD = 15.4); Min = 19, Max = 90	39	Mean = 47.4 (SD = 19.2); Min = 19, Max = 84	$t_{185,38} = 2.4$; $P = .012$

Abbreviations: y, years; SD, standard deviation; Min, minimum; Max, maximum.

Table 2
Concordance of sleep disorder diagnoses based on the Diagnostic Interview for Sleep Patterns and Disorders interview compared with expert clinician (Referent).

Sleep disorder	Clinician diagnosis	Interview diagnosis	Sensitivity		Specificity		Total classification accuracy		McNemar χ^2 test		κ		AUC
	n (%) (Referent)	*n* (%)	%	SE	%	SE	%	SE	Stat	*P* value	κ	SE	
Delayed sleep phase	30 (13.3)	33 (14.7)	0.63	0.09	0.93	0.02	0.89	0.02	0.36	.55	0.54	0.08	0.781
Hypersomnia	69 (30.7)	72 (32.0)	0.61	0.06	0.81	0.03	0.75	0.03	0.16	.69	0.41	0.07	0.708
Insomnia	98 (43.6)	156 (69.3)	0.87	0.03	0.44	0.04	0.63	0.03	40.1	.00	0.29	0.05	0.654
Narcolepsy with cataplexy	28 (12.4)	19 (8.4)	0.68	0.09	1.00	0.00	0.96	0.01	9.00	.00	0.79	0.07	0.839
Period limb movement disorder	33 (14.7)	46 (20.4)	0.67	0.08	0.88	0.02	0.84	0.02	4.83	.03	0.47	0.08	0.771
Restless legs syndrome	14 (6.2)	65 (28.9)	0.79	0.11	0.74	0.03	0.75	0.03	45.6	.00	0.20	0.06	0.765
Rapid eye movement sleep behavior disorder	8 (3.6)	13 (5.8)	0.50	0.18	0.96	0.01	0.94	0.02	1.92	.17	0.35	0.14	0.729
Sleep apnea	89 (39.6)	115 (51.1)	0.75	0.05	0.65	0.04	0.69	0.03	9.66	.00	0.38	0.06	0.700

Abbreviations: SE, standard error; AUC, area under the receiver operating characteristic curve.

sleep disorders according to ICSD-2 adult criteria [31]. NC, RLS, RBD, and delayed sleep phase disorder criteria were based on the ICSD-2 criteria. The ICSD-2 criteria for insomnia required symptom present for at least one month. OSA also was based on the ICSD-2 criteria including PSG findings with more than five apnea or hypopnea events per hour. At the Chevy Chase Center, hypopneas were defined following the American Academy of Sleep Medicine alternative criteria, which requires at least 50% reduction in nasal pressure signal excursions and associated $\geqslant$3% desaturation or arousal. At the Stanford Center, hypopneas were defined as a 30% (or greater) reduction in flow lasting 10 s or longer accompanied by a 3% (or greater) desaturation in the oxygen levels or an arousal, which is a slight modification of the American Academy of Sleep Medicine alternate criteria. Esophageal manometry was not used. The presence of periodic leg movements was defined as a periodic leg movement index score of $\geqslant$15 per hour, without any need for daytime symptoms. Hypersomnia was defined as complaints of excessive daytime sleepiness symptoms occurring almost daily for at least 3 months and an MSLT documenting shortened sleep onset latency (<8 min). Patients with a diagnosis of narcolepsy were not considered as hypersomnia cases. Different subtypes of hypersomnia, such as idiopathic hypersomnia with long sleep duration, were not further classified.

2.3. Procedure

Patients referred to the sleep centers were invited to participate in either a face-to-face or telephone interview about their sleep patterns and disorders. Study participants were referred to the NIH research coordinator who was blind to the referral diagnosis at the sleep center. To maintain the blindness of the interviewer, study participants were instructed to refrain from revealing their diagnosis or treatment prior to administration of the interview. The study was approved by the Combined Neurosciences Review Committee of the NIH and the Human Investigation Committee of Stanford University.

2.4. Analysis

Diagnostic classifications based on clinician (treated as referent, "true") and from those made by computer algorithms based on sleep interview were examined using McNemar χ^2 tests. McNemar tests are explicitly designed for paired comparison of dichotomies. Breslow–Day tests were used to test for homogeneity between the sleep centers. Convergent and discriminant validity was assessed by examining the concordance between cases identified with the clinician and sleep interview based on 2×2 contingency tables. Standard descriptive statistics used to analyze such tables were calculated: (1) sensitivity (i.e., the proportion of true cases [by clinician diagnosis] that screened positively [by sleep interview]); (2) specificity (i.e., the proportion of true noncases that screened negatively); (3) positive predictive value (i.e., the proportion of screened cases that were true cases); and (4) negative predictive value (i.e., the proportion of screened noncases that were true noncases). Three summary measures of overall individual-level concordance, total classification accuracy, Cohen κ coefficient [43], and area under the receiver-operating characteristic curve (AUC) [44] were calculated. Total classification accuracy tends to be biased towards high levels of agreement in situations in which the prevalence of a disorder is low [45,46]. The κ coefficient is the most widely used measure of concordance in validity studies of psychiatric disorders, but it has been criticized as it is dependent on prevalence and consequently often is low in situations in which there appears to be high agreement between low-prevalence measures [45,47,48]. The AUC, our main focus, was calculated as a measure of the extent to which the screened cases predicted a true diagnosis. This prevalence-free statistic can be interpreted as the probability that a randomly selected true case and a randomly selected true noncase would be correctly distinguished based on the sleep questionnaire interview. If the screen diagnosis is completely unrelated to the true disorder, the expected value of the AUC is 0.5, while if the screen diagnosis is a perfect predictor of a true diagnosis, the AUC will be 1.0. Concordance can be described as almost perfect (AUC, 0.91–1.0), substantial (AUC, 0.81–0.9), moderate (AUC, 0.71–0.8), fair (AUC, 0.61–0.7), and slight (AUC, 0.51–0.6).

3. Results

The clinical validity statistics are presented in Table 2. The rates of disorders based on the sleep interview diagnosis were significantly higher than the clinician diagnosis for four of the eight disorders (insomnia, PLMD, RLS, and OSA). The largest difference involved RLS (28.9% sleep interview vs 6.2% clinician diagnosis). In contrast, the rate of NC based on the sleep interview was significantly lower than that of the clinician (8.4% vs 12.4%). The estimates were comparable for delayed sleep phase disorder, hypersomnia, and RBD. Homogeneity tests showed that there were marginal differences in the rates of delayed sleep phase disorder diagnosis across the two sleep centers ($P = .050$), whereas there was no difference in the other seven sleep disorders.

Overall, the sleep interview tended to classify more subjects as cases than those who were classified by the clinician diagnosis. The interview diagnoses had fair to substantial concordance with clinician diagnoses with AUCs ranging from 0.65 to 0.84. The sleep interview diagnosis had substantial agreement with clinician diagnosis for NC (AUC, 0.84); moderate agreement for delayed sleep phase disorder, hypersomnia, PLMD, RLS, and OSA (AUC, 0.71–0.78); and only fair agreement for insomnia (AUC, 0.65). The average number of sleep disorders was 2.3 by sleep interview, compared to 1.6 by clinician diagnosis. The DISP yielded two or more sleep disorders in 68.4% of the sample, whereas the clinician records only identified two or more sleep disorders among 44.0% of the patients.

Participants classified by a clinician as having a sleep disorder were moderately well-detected as cases by the sleep interview (sensitivity ranged from 0.50 for RBD to 0.87 for insomnia). The majority of clinician-diagnosed noncases also were classified by the sleep interview as noncases, with specificity estimates ranging from 0.44 for insomnia to 1.00 for NC. Substantial specificity (>0.8) was found for five of eight sleep disorders (delayed sleep phase disorder, hypersomnia, NC, PLMD, and RBD). The sleep interview had increased sensitivity over clinician diagnosis in detecting six of the eight sleep disorders, particularly RLS and hypersomnia, which were the index referral diagnoses in our study. The negative predictive values for all sleep disorders exceeded 0.80 with five greater than 0.90, indicating that the sleep interview correctly classified clinician-identified noncases. However, the positive predictive values for most sleep disorders except for NC, were less than 0.60.

4. Discussion

Our findings provide initial evidence regarding the validity of the DISP, the first structured interview of the full range of sleep disorders that can be administered by trained lay interviewers. The total classification accuracy was more than 65% for each sleep disorder and more than 80% for delayed sleep phase disorder, NC, PLMD, and RBD. In addition, there was substantial specificity (>0.8) for five of the eight sleep disorders assessed in the DISP. Therefore, the DISP provides an efficient (average administration time of 30 min) and relatively inexpensive method of detecting specific sleep disorders that may be used in both clinical and community settings. The DISP also identified secondary nonreferral sleep disorders that were not identified in the clinical settings.

The high classification accuracy for delayed sleep phase disorder and NC is not surprising, considering that these diagnoses are primarily based on symptoms. These conditions rarely are missed by sleep experts, unlike in RLS and insomnia, which often may be overlooked by clinicians. This finding is particularly true when the sleep disorders are secondary to another referral diagnosis, as demonstrated in our study. The accuracy of the DISP also is comparable to that of previous sleep disorder measures; however, such direct comparisons are inherently limited by variation in study design, sample size, and selection. The DISP has slightly lower classification accuracy (range of κ coefficients, 0.20–0.79) than two previous Diagnostic Interviews that also cover a range of sleep disorders (range of κ coefficients, 0.38–1.0) [19,49]. However, the validity of these interviews was based on a limited number of cases, thereby limiting cross-instrument comparability. Regarding measures of specific sleep disorders, the DISP was comparable to interviews that did not include actual laboratory measures as validation criteria. The somewhat lower reliability of the DISP to detect insomnia compared to that of two widely used sleep questionnaires, the Pittsburgh Sleep Quality Index [25] and the Insomnia Severity Inventory [50], is likely to be attributable to different time periods of assessment; the latter covers the past 2–4 weeks, whereas our study measured the validity of lifetime insomnia diagnosis. The DISP also was less reliable than the Johns Hopkins Telephone Diagnostic Interview for RLS [21], which is administrated by clinicians and assesses different diagnostic criteria than the ICSD-2. The AUC of 0.70 for OSA is lower than questionnaires such as the Multivariable Apnea Prediction questionnaire [23]. However, the performance of the DISP (AUC, 0.700) in differentiating OSA was close to that of the Multivariable Apnea Prediction index (AUC, 0.695), without body mass index. Sensitivity was similar (0.75 vs 0.77), but specificity was lower (0.65 vs. 0.89) in predicting ICSD-2 criteria in the DISP when compared to the Berlin Questionnaire

in predicting the apnea–hypopnea index of ⩾5 events per hour [24]. However, this comparison is less relevant, as the latter questionnaire does not assess the ICSD-2 criteria that were the target of the DISP.

The DISP also identified several secondary conditions, particularly insomnia and RLS, which were not detected by the expert sleep clinicians. The average number of sleep disorders detected by the DISP was 2.3 compared to 1.6 by clinical record diagnosis. It is plausible that clinicians focus on the primary sleep disorder without assessing full criteria for other secondary sleep problems, particularly in the context of referral from another clinician, specifically for evaluation of a particular sleep disorder. The possible underestimation of some sleep disorders in clinical settings also was evident in our analyses of the prevalence rates of physician-diagnosed insomnia and RLS in a nationally representative sample that yielded substantially lower prevalence estimates of these conditions than those assessed by questionnaire or clinical interview [4,51]. This finding implies that estimates of the magnitude of sleep disorders that primarily reflect subjective reporting of symptoms based solely on clinical diagnosis such as insomnia and RLS may be underestimates of the true population prevalence of these conditions. On the other hand, the sleep interview could have yielded false positive rates of these disorders. Although increased false positive rates are unlikely to apply to insomnia, which has fairly clear-cut operational criteria that was further validated in subsequent follow-up interviews of a substantial subset of our sample, it is possible that interviewer estimates of RLS could have been inflated based on their inability to detect mimics of RLS and trained clinicians [22].

Strengths of our study and interview included (1) our design, which compared ICSD-2–defined sleep disorders generated by blind independent assessment of sleep disorders in a carefully diagnosed sample of individuals from nationally recognized sleep experts who considered information on the clinical history, clinical examination, nocturnal PSG, and MSLT, when relevant; (2) inclusion of a wide range of sleep disorders based on ICSD-2 criteria; (3) a structured format and scoring algorithms which strongly reduced interrater reliability induced by subjective judgment and diagnostic coding errors; (4) specific modules which could be used alone due to their independence in the DISP; (5) combined information on sleep disorders and sleep patterns (i.e., sleep schedule, daytime napping, sleep regularity, circadian preference), which provided a context of sleep patterns that could be used to assess sleep disturbances and correlates of sleep disorders collected in subsequent sections of the interview; and (6) inclusion of several key correlates of clinical phenomenology, such as the number of episodes, course, impairment, and treatment history, which facilitated establishment of lifetime patterns that were highly relevant for disorders and that may wax and wane across the life course. The latter feature is particularly important for studies in genetic epidemiology for which lifetime diagnosis is required to evaluate patterns of heritability in families. The use of questionnaires that assess current symptoms generate substantial false negatives, depending on the extent to which the sleep disorder is chronic vs episodic such as in insomnia. For example, estimates of the heritability of insomnia are based on studies that have solely assessed insomnia symptoms [52,53] and that have been shown to yield underestimates of the heritability of insomnia disorders [54].

Limitations of the our study and interview also should be considered in interpreting our findings. First, the validity of the DISP was diminished by the design, which excluded incorporation of lifetime as opposed to current diagnosis, test results, and treatment to maximize blindness of the interviewers. Inclusion of such information could strongly enhance its clinical utility. Employment of symptom-level data also might help to improve the validity of the DISP. Second, the relatively small numbers of participants with some of the less common sleep disorders reduced the power to quantify agreement; however, the number of cases for most sleep disorders were larger than previous validation studies for existing sleep interviews [19,49]. In addition to the differential numbers, the actual validity of the DISP was variable across different sleep disorders. For example, the total classification accuracy was less than 70% for insomnia and OSA. Third, the lifetime diagnoses were determined through retrospective report, which may imply recall bias. Fourth, the application of categorical diagnosis rather than dimensional ratings of symptoms derived from the interview may have led to lower agreement [18]. Fifth, the samples identified in nationally recognized sleep centers were not representative of those from the greater community in education (75% with a college or above education level) or clinical phenomenology. Therefore, we have devoted substantial effort to enhancing the generalizability of the findings from our study to a large community-based local sample, which may generate further adaptation of our interview. We have administered the DISP to more than 1000 participants in a community-based family study of health and behavior underway in our program to document its feasibility and clinical reliability. However, formal validation of community-identified cases requires full evaluation including PSG, which is beyond the scope of resources of this program. Furthermore, we are studying the reliability and validity of a Chinese version of the DISP in a broader sample of patients with sleep disorders in Hong Kong. Sixth, although interrater reliability was informally established during the pilot phase of our study, we did not calculate formal reliability statistics. Moreover, because of the cross-sectional nature of this validation study, we did not assess the test–retest reliability of the DISP. Ongoing follow-up of our sample will enable us to test the stability of reported sleep disorders over time. Finally, we did not compare expert clinician with interviewer diagnoses on the DISP. However, our extensive experience with such structured diagnoses in psychiatry has yielded few differences between experienced clinicians and well-trained and monitored lay interviewers due to the highly structured format of the interview.

5. Conclusion

To our knowledge, our interview is the first structured interview which can be administered by a trained lay interviewer to screen for the full range of sleep disorders and their clinical impact. It provides a time-efficient (average administration time of 30 min) and relatively inexpensive method of detecting specific sleep disorders in both clinical and community settings. Although our validation study was conducted with patients from nationally recognized sleep centers, our administration of the DISP to more than 1000 individuals in a community-based family study of comorbidity of mood disorders and cardiovascular disease also demonstrates its feasibility in community settings.

The detection of secondary primarily syndromic conditions in the DISP, which were not identified in the sleep center records, also highlights the value of administration of structured Diagnostic Interviews for sleep disorders to capture the full range of sleep disorders in clinical settings. Such secondary conditions may be highly relevant to the treatment, course, and prevention of the consequences of sleep disorders. Ongoing studies of the DISP in community settings will provide additional information on its public health utility and cross-cultural validity.

Conflict of interest

The ICMJE Uniform Disclosure Form for Potential Conflicts of Interest associated with this article can be viewed by clicking on the following link: http://dx.doi.org/10.1016/j.sleep.2013.10.011.

Acknowledgements

This work was supported by the Intramural Research Program of the National Institute of Mental Health (Z01 MH002804). The views and opinions expressed in this article are those of the authors and should not be construed to represent the views of any of the sponsoring organizations, agencies, or US Government.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.sleep.2013.10.011.

References

[1] Buysse DJ. Insomnia. JAMA 2013;309:706–16.
[2] Jansson C, Alexanderson K, Kecklund G, Akerstedt T. Clinically diagnosed insomnia and risk of all-cause and diagnosis-specific sickness absence: a nationwide Swedish prospective cohort study. Scand J Public Health 2013;41:712–21.
[3] Bin YS, Marshall NS, Glozier N. The burden of insomnia on individual function and healthcare consumption in Australia. Aust N Z J Public Health 2012;36:462–8.
[4] Allen RP, Bharmal M, Calloway M. Prevalence and disease burden of primary restless legs syndrome: results of a general population survey in the United States. Mov Disord 2011;26:114–20.
[5] Budhiraja P, Budhiraja R, Goodwin JL, Allen RP, Newman AB, Koo BB, et al. Incidence of restless legs syndrome and its correlates. J Clin Sleep Med 2012;8:119–24.
[6] Young T, Peppard PE, Gottlieb DJ. Epidemiology of obstructive sleep apnea: a population health perspective. Am J Respir Crit Care Med 2002;165:1217–39.
[7] Leger D, Poursain B. An international survey of insomnia: under-recognition and under-treatment of a polysymptomatic condition. Curr Med Res Opin 2005;21:1785–92.
[8] Walsh JK. Clinical and socioeconomic correlates of insomnia. J Clin Psychiatry 2004;65(Suppl. 8):S13–9.
[9] Novak M, Mucsi I, Shapiro CM, Rethelyi J, Kopp MS. Increased utilization of health services by insomniacs – An epidemiological perspective. J Psychosom Res 2004;56:527–36.
[10] Allen RP, Walters AS, Montplaisir J, Hening W, Myers A, Bell TJ, et al. Restless legs syndrome prevalence and impact: REST general population study. Arch Intern Med 2005;165:1286–92.
[11] Young T, Finn L, Peppard PE, Szklo-Coxe M, Austin D, Nieto FJ, et al. Sleep disordered breathing and mortality: eighteen-year follow-up of the Wisconsin sleep cohort. Sleep 2008;31:1071–8.
[12] American Academy of Sleep Medicine. International classification of sleep disorders: diagnostic and coding manual (ICSD). Rochester, Minnesota: American Academy of Sleep Medicine; 1990.
[13] Ip MS, Lam B, Lauder IJ, Tsang KW, Chung KF, Mok YW, et al. A community study of sleep-disordered breathing in middle-aged Chinese men in Hong Kong. Chest 2001;119:62–9.
[14] Leger D, Guilleminault C, Dreyfus JP, Delahaye C, Paillard M. Prevalence of insomnia in a survey of 12,778 adults in France. J Sleep Res 2000;9:35–42.
[15] Karacan I, Thornby JI, Anch M, Holzer CE, Warheit GJ, Schwab JJ, et al. Prevalence of sleep disturbance in a primarily urban Florida County. Soc Sci Med 1976;10:239–44.
[16] Ohayon MM, Roth T. Prevalence of restless legs syndrome and periodic limb movement disorder in the general population. J Psychosom Res 2002;53:547–54.
[17] Berger K, Luedemann J, Trenkwalder C, John U, Kessler C. Sex and the risk of restless legs syndrome in the general population. Arch Intern Med 2004;164:196–202.
[18] Kessler RC, Coulouvrat C, Hajak G, Lakoma MD, Roth T, Sampson N, et al. Reliability and validity of the brief insomnia questionnaire in the America insomnia survey. Sleep 2010;33:1539–49.
[19] Ohayon MM, Guilleminault C, Zulley J, Palombini L, Raab H. Validation of the sleep-EVAL system against clinical assessments of sleep disorders and polysomnographic data. Sleep 1999;22:925–30.
[20] Abetz L, Arbuckle R, Allen RP, Mavraki E, Kirsch J. The reliability, validity and responsiveness of the Restless Legs Syndrome Quality of Life questionnaire (RLSQoL) in a trial population. Health Qual Life Outcomes 2005;3:79.
[21] Hening WA, Allen RP, Thanner S, Washburn T, Heckler D, Walters AS, et al. The Johns Hopkins telephone diagnostic interview for the restless legs syndrome: preliminary investigation for validation in a multi-center patient and control population. Sleep Med 2003;4:137–41.
[22] Hening WA, Allen RP, Washburn M, Lesage SR, Earley CJ. The four diagnostic criteria for restless legs syndrome are unable to exclude confounding conditions (“mimics”). Sleep Med 2009;10:976–81.
[23] Maislin G, Pack AI, Kribbs NB, Smith PL, Schwartz AR, Kline LR, et al. A survey screen for prediction of apnea. Sleep 1995;18:158–66.
[24] Netzer NC, Stoohs RA, Netzer CM, Clark K, Strohl KP. Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome. Ann Intern Med 1999;131:485–91.
[25] Buysse DJ, Reynolds 3rd CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989;28:193–213.
[26] Kim K, Uchiyama M, Okawa M, Liu X, Ogihara R. An epidemiological study of insomnia among the Japanese general population. Sleep 2000;23:41–7.
[27] Hajak G. Epidemiology of severe insomnia and its consequences in Germany. Eur Arch Psychiatry Clin Neurosci 2001;251:49–56.
[28] Chevalier H, Los F, Boichut D, Bianchi M, Nutt DJ, Hajak G. Evaluation of severe insomnia in the general population: results of a European multinational survey. J Psychopharmacol 1999;13(4 Suppl. 1):S21–4.
[29] Mellinger GD, Balter MB, Uhlenhuth EH. Insomnia and its treatment. Prevalence and correlates. Arch Gen Psychiatry 1985;42:225–32.
[30] Ohayon MM. Epidemiology of insomnia: what we know and what we still need to learn. Sleep Med Rev 2002;6:97–111.
[31] Breslau N, Roth T, Rosenthal L, Andreski P. Sleep disturbance and psychiatric disorders: a longitudinal epidemiological study of young adults. Biol Psychiatry 1996;39:411–8.
[32] Bixler EO, Kales A, Soldatos CR, Kales JD, Healey S. Prevalence of sleep disorders in the Los Angeles metropolitan area. Am J Psychiatry 1979;136:1257–62.
[33] Lavigne GJ, Montplaisir JY. Restless legs syndrome and sleep bruxism: prevalence and association among Canadians. Sleep 1994;17:739–43.
[34] Cho YW, Shin WC, Yun CH, Hong SB, Kim JH, Allen RP, et al. Epidemiology of restless legs syndrome in Korean adults. Sleep 2008;31:219–23.
[35] Hogl B, Kiechl S, Willeit J, Saletu M, Frauscher B, Seppi K, et al. Restless legs syndrome: a community-based study of prevalence, severity, and risk factors. Neurology 2005;64:1920–4.
[36] Phillips B, Young T, Finn L, Asher K, Hening WA, Purvis C. Epidemiology of restless legs symptoms in adults. Arch Intern Med 2000;160:2137–41.
[37] Young T, Shahar E, Nieto FJ, Redline S, Newman AB, Gottlieb DJ, et al. Predictors of sleep-disordered breathing in community-dwelling adults: the Sleep Heart Health Study. Arch Intern Med 2002;162:893–900.
[38] Cirignotta F, D'Alessandro R, Partinen M, Zucconi M, Cristina E, Gerardi R, et al. Prevalence of every night snoring and obstructive sleep apnoeas among 30–69-year-old men in Bologna, Italy. Acta Neurol Scand 1989;79:366–72.
[39] Duran J, Esnaola S, Rubio R, Iztueta A. Obstructive sleep apnea–hypopnea and related clinical features in a population-based sample of subjects aged 30–70 yr. Am J Respir Crit Care Med 2001;163(3, pt. 1):685–9.
[40] Tishler PV, Larkin EK, Schluchter MD, Redline S. Incidence of sleep-disordered breathing in an urban adult population: the relative importance of risk factors in the development of sleep-disordered breathing. JAMA 2003;289:2230–7.
[41] Bixler EO, Vgontzas AN, Lin HM, Calhoun SL, Vela-Bueno A, Kales A. Excessive daytime sleepiness in a general population sample: the role of sleep apnea, age, obesity, diabetes, and depression. J Clin Endocrinol Metab 2005;90:4510–5.
[42] PhenX Toolkit. https://www.phenxtoolkit.org/index.php?pageLink=browse.protocoldetails&id=121001. Published May 12; 2010. Accessed December 08, 2013.
[43] Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960;20:37–46.
[44] Hanley JA, Mcneil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29–36.
[45] Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol 1993;46:423–9.
[46] Kraemer HC, Morgan GA, Leech NL, Gliner JA, Vaske JJ, Harmon RJ. Measures of clinical significance. J Am Acad Child Adolesc Psychiatry 2003;42:1524–9.
[47] Cook RJ, Farewell VT. Conditional inference for subject-specific and marginal agreement: two families of agreement measures. Can J Stat 1995;26:333–44.
[48] Cicchetti DV, Feinstein AR. High agreement but low kappa II: resolving the paradoxes. J Clin Epidemiol 1990;43:551–8.
[49] Schramm E, Hohagen F, Grasshoff U, Riemann D, Hajak G, Weess HG, et al. Test-retest reliability and validity of the Structured Interview for Sleep Disorders According to DSM-III-R. Am J Psychiatry 1993;150:867–72.
[50] Bastien CH, Vallieres A, Morin CM. Validation of the insomnia severity index as an outcome measure for insomnia research. Sleep Med 2001;2:297–307.
[51] Roth T, Jaeger S, Jin R, Kalsekar A, Stang PE, Kessler RC. Sleep problems, comorbid mental disorders, and role functioning in the national comorbidity survey replication. Biol Psychiatry 2006;60:1364–71.
[52] Gehrman PR, Meltzer LJ, Moore M, Pack AI, Perlis ML, Eaves LJ, et al. Heritability of insomnia symptoms in youth and their relationship to depression and anxiety. Sleep 2011;34:1641–6.
[53] Drake CL, Friedman NP, Wright Jr KP, Roth T. Sleep reactivity and insomnia: genetic and environmental influences. Sleep 2011;34:1179–88.
[54] Wing YK, Zhang J, Lam SP, Li SX, Tang NL, Lai KY, et al. Familial aggregation and heritability of insomnia in a community-based study. Sleep Med 2012;13:985–90.